Appendix H
Data Validation

Appendix H, Data Validation

Table of Contents

Attachments

1.0 Introduction

Data validation is performed to determine how well project data meet project acceptance criteria. If potentially severe data quality problems are identified in a data review, then project management should consider a full-scale validation effort.

1.1 Purpose

This document provides guidance for the technical validation of analytical data generated and used in support of the Navy's Installation restoration (IR) program.

The Engineering Field Division/Engineering Field Activity (EFD/EFA), as the project manager, shall establish the required frequency and level of effort for data validation in project planning documents and should define the process through which the specific data intended for validation will be selected. It should be noted that the direction given in this section is typically presented using the term "should." The EFD/EFA shall determine if the information and guidance presented in this section shall be applied more stringently (i.e. "should" implemented as "shall").

1.2 Objectives

This document provides guidance for the scope, context, and approach for data validation. This document is not intended to serve as a standard operating procedure (SOP) for data validation activities.

1.3 Scope

This document describes the general elements of technical reviews of data generated using nonprescriptive methods such as the Environmental Protection Agency (EPA) SW-846 and highly prescriptive methods such as the EPA Contract Lab Program (CLP). It is applicable to reviews of chemical data generated using published reference methods. Although the data review process is generally applicable, the user is cautioned that the technical details in this document are not universally applicable to data generated using all published methods (e.g., they would not be applicable to bioassay, radiochemical or geological testing).

2.0 Data Review and Data Validation

Data review and data validation are not adequately defined in most procedures or guidance documents. For purposes of this document, data review is defined as a systematic approach for the review of laboratory data. Data validation is a thorough assessment of data and supporting QC documentation without making any assumption to the quality of the data provided.

2.1 Data Review

In a summary or low level review only the sample results and limited project documentation are typically reviewed. Summary or low-level reviews are best suited to cases in which some project data has been subjected to a high level or full-scale validation.

Typically, laboratory personnel and end users perform data reviews as a quality assurance/quality control (QA/QC) measure. It is the responsibility of end users to review 100 percent of laboratory data for completeness. This type of review is commonly referred to as "summary level" review. In summary level reviews, the following elements should be examined:

- Completeness: Determine if:
  - All requested analytes accounted for
  - All Project Data Quality Objectives (DQOs) or target/action levels met
  - Results correlate with historic data
- Holding times: Are they within limits
- Chain of custody: Is documentation complete and accurate
- Method and reporting limits: Are they within the scope of project DQOs
- Dilution factors/concentration units: Are they correct as reported
- Preparation/analysis methods: Were those identified on the report appropriate for the project
- Matrix spike results (if provided): Were they within specification
- Surrogate recoveries (if provided) within specification

The results of a summary level review may reveal inaccuracies or errors in the data that may require a more thorough assessment, such as data validation.

2.2 Data Validation

In a full level data validation, validators review and evaluate reported data, raw data, supporting information, and project documentation to make a determination as to whether the reported data are of sufficient quality to satisfy project objectives.

In many cases, project plans and management reviews do not specify the elements that must be reviewed for data validation. If specific project or program guidance (i.e., the

CLP) is not available to determine the elements necessary for data validation, then the guidance in the following sections may be applied.

2.2.1 Validation Scope

The appropriate scope for project data validation should be determined in consideration of project DQOs and established in project planning documents such as a QA Plan. Planning documents should specify:

- Which data set(s) will be subject to validation by sample type, location, or sampling period as appropriate
- The frequency or percentage of data to be validated
- The level or degree of validation required and the specific laboratory documents required to accomplish the validation
- The source or reference documents used to determine applicable technical performance (i.e., to qualify or "flag" the data)

The overall scope of a project's data validation effort may be relatively large for data that is critical to providing input for decisions involving high risk or low tolerance for risk. Conversely, limited or no validation may be required for routine project data.

Data validation may be scoped as a full-scale effort or limited to only a summary level review without data validation.

2.2.1.1 Navy QA Guidance

The Navy Installation Restoration Chemical Data Quality Manual (IR CDQM) provides or references QC requirements and criteria that must be adopted and implemented by a laboratory in the absence of project-specific instructions. A copy of the guidance document must be provided to the data validators for reference. However, data validation is not intended to assess a laboratory's compliance with the IR CDQM.

2.2.2 Validation Levels

All aspects of the data are reviewed and appropriate data "flags" assigned. The Navy understands that a consensus among agencies does not exist for the degree of documentation review required to "validate" data. Therefore, it is important that project plans specifically outline the areas of lab documentation that must be reviewed prior to validation of data.

Some agencies, such as EPA, create "levels" or "tiers" of review for data validation. An upper level or tier review may require extensive documentation and research including calibrations, standards traceability, contract review, statement of work review (SOW), and on-site lab audits.

As an end user of data, it is important to remember that the amount of documentation required to perform data reviews (level/tier I, II, III, etc.) is not consistent among agencies. Level II validation, for example, may not have the same meaning in various EPA regions or military components. Data validation reports must include as references any documents(s) that were used to determine the degree or level of validation required.

If project specific plans or responsible regulatory authority do not specify the required criteria for data validation, then the following will apply.

2.2.2.1 CLP

For data under the CLP, use as references:

- Applicable EPA Region Quality Assurance Project Plan Guidance
- EPA Regional Data Validation Functional Guidelines for Evaluating Environmental Analyses
- USEPA Contract Laboratory Program National Functional Guidelines for Organic Data Review
- USEPA Contract Laboratory Program National Functional Guidelines for Inorganic Data Review

For data generated outside the Contract Lab Program, the guidance provided in section 2.2.2.2 will apply.

2.2.2.2 Non CLP

After reviewing all elements of a data set, the validator will use professional judgement to generate an overall summary of the technical quality of the data. Data may be qualified or "flagged" based on the elements presented in the tables included in Attachment (1) to this appendix, or for other reasons as documented by the validators. Data must not be qualified if only a summary level review of the data set was performed. Final qualifiers for individual data points that have previously been flagged with multiple individual qualifiers may be increased in severity (e.g., a point which has been flagged with multiple "J"s on the basis of several different quality elements may be downgraded to an "R" flag). The validator must document any suspected biases in the data set.

The end user must determine the impact of all suspected biases or qualified data based on intended project use.

2.3 Frequency of Review and Validation

As stated in Section 2.1, end users must review 100 percent of the data for which they are responsible. For a given project a certain percentage of the data may require a more thorough assessment, or data validation. The frequency of validation should be

determined based on consideration of project DQOs. Each project's planning documents should establish the required data validation frequency, define the process through which specific data is selected for validation, and the level of detail in documentation required to validate the data.

3.0 Validator Qualifications

3.1 Education

The individuals who provide data validation services must have technically appropriate credentials that are commensurate with their responsibilities. The individuals who perform or review data validation must have a minimum of a Bachelor of Science (BS) or Bachelor of Arts (BA) in chemistry or a physical science. Validators who do not meet these requirements should provide documented evidence that demonstrates that they possess the disciplinary expertise, experience and theoretical knowledge necessary to validate data.

3.2 Knowledge and Experience

Each individual who provides data validation services must have a minimum of 2 years of professional bench level experience beyond a baccalaureate degree that is commensurate with their method specific responsibilities for data validation. For example:

- To validate data from volatile organic methods, an individual must have performed GC-MS analyses for the determination of trace level volatile organic contaminants.
- To validate data from semi-volatile organic methods, an individual must have performed GC-MS or HPLC analyses for the determination of trace level semi-volatile organic contaminants.
- To validate data from Pesticide/PCB methods, an individual must have performed GC analyses for the determination of trace level organics.
- To validate data from Dioxin methods, an individual must have performed GC-MS analyses and must have experience using high-resolution mass spectroscopy techniques.
- To validate data from metals methods, an individual must have performed analyses for the determination of trace metals using ICP, ICP-MS, AA, or GFAA.

To validate data from classical methods (e.g., CRVI, CN, ion chromatography, gravimetric, etc.) or radiochemical methods (gross alpha, beta, gamma, etc.) an individual must have performed analyses using the referenced methods.

Appendix H
Attachment 1
Objective, Subjective, and Supporting Data Elements Reviewed During Data Validation

1.0 Objective Data Elements Reviewed During Data Validation

The initial step in the data validation process is the review and evaluation of objective data elements. The objective data elements addressed in this section are independent of sample matrix, and provide objective, quantitative information regarding performance of the preparative and analytical methods and instrumentation (if applicable) during the measurement process. Compliance with individual method, project, or Navy acceptance criteria must be evaluated, as appropriate. Data associated with unacceptable QC may be of extremely limited use and must be carefully assessed and qualified if the data are not to be rejected. Table H-1 summarizes the objective data elements that must be reviewed during the validation process:

Table H-1.

Validation Element	Criteria and specifications to be assessed during validation
Initial Calibration	- number of standards used - range of calibration - algorithm used - samples analyzed and reported within calibration range
Initial Calibration Verification (ICV)	- independent, or second source standard - concentration - percent recovery - position in the analytical run sequence(s)
Initial Calibration Blank (ICB)	- composition of the blank - analytical result(s) - position in the run sequence(s)
Continuing Calibration Verification (CCV)	- concentration - percent recovery - position in the analytical run sequence(s) - frequency
Laboratory Control Sample (LCS)	- composition (matrix) - concentration - percent recovery - trends in LCS recovery (if possible)
Laboratory Control Sample Duplicate (LCSD)	- evaluation criteria, as for LCS - performance of LCSD appropriate - batch precision

Interference Check Standard (ICP, ICP/MS)	- composition - concentration - percent recovery - position(s) in the analytical run sequence(s)
Method Blank (MB)	- detection of target analytes - concentration of target analytes - percent recovery of compounds added (e.g., surrogates)
Instrument Blanks	- detection and concentration of target analytes - percent recovery of any compounds added (e.g., surrogates)
Process Blanks (e.g., trip blanks, holding blanks, and rinsate blanks)	- detection and concentration of target analytes - percent recovery of any compounds added (e.g., surrogates)
GC/MS Tunes	- compound used - amount analyzed - introduction technique - instrument operating parameters - spectrum generation procedure
GC Degradation Check	- compounds used - standard concentrations - algorithm for breakdown calculation - compliance with criteria
GC and LC Retention Time Windows	- number of standards analyzed - temporal spacing of analyses - algorithm for calculation of window size - procedure for centering windows - frequency of recentering
HRMS Resolution and Mass Accuracy (Dioxins)	- resolution - mass accuracy
Dioxin GC Column Performance Check	- resolution of 2,3,7,8-TCDD - retention times of analytes
Analytical Wavelength (ICP, spectrophotometric analysis)	- analytical wavelengths used - consistency with QC and method performance data
Method of Standard Addition (GFAA)	- spike concentrations - number of concentration levels - algorithm for calculation of sample concentration

High Calibration Standard (ICP)	- position in analytical run sequence - acceptance with criteria
Gel Permeation Chromatography (GPC)	- analytical results for the GPC blank - calibration check
Linear Range	- samples analyzed within linear range - reasonableness of linear ranges determined
Calculations	- confirmation of manual calculations
Samples	- assessment of results (i.e., detection, qualitative identification, and quantitation) with reference to all objective validation elements

2.0 Subjective Data Elements and Criteria

The second step in the data validation process is the review and evaluation of subjective data elements. The effect of these review elements on the integrity and usability of the data set must be assessed using professional judgment. Although some elements are assigned numerical values and acceptance criteria, the relationship of the numerical value to data validity, acceptability, accuracy, and precision cannot be precisely and predictably determined. The impact of these subjective data elements on data validity, usability, and defensibility must be assessed, and data qualified as warranted in consideration of project objectives. Table H-2 summarizes the subjective data elements that must be reviewed as part of the validation process:

Table H-2. Subjective Data Elements Reviewed During Data Validation

Validation Element	Criteria and specifications to be assessed during validation
Matrix Spike/Matrix Spike Duplicate (MS/MSD)	- concentration - percent recovery - precision - field duplicate precision
Duplicates (Laboratory and Field)	- precision - sample heterogeneity - subsample heterogeneity
Hold-Time	- verify to preparation and analysis dates, as appropriate
Serial Dilutions (ICP)	- appropriate sample diluted - percent difference between diluted and undiluted sample result
Post Digestion Spike (PDS)	- spike concentration - percent recovery
Surrogates	- surrogates used - calibration and quantitation procedures - concentrations - percent recovery
Internal Standard Responses	- internal standards used - concentrations in standards and extracts/digestates - instrument responses
Organic Internal Standard Retention Times	- retention times
GC and LC Confirmation Analyses	- procedures for confirmation analyses (e.g., initial calibrations, calibration verifications, etc.) - procedures for combining results from two analyses
Coeluting Compounds in GC and LC Analyses	- procedures for treatment of coeluting compounds

Qualitative Identification of GC/LC Target Compounds	- procedures for use of retention time windows or pattern matching
Qualitative Identification of GC/MS Target Compounds and TICs	- relative retention times (target compounds only) and spectra of reported compounds - closely eluting compounds with similar spectra
Qualitative Identification of Dioxins	- relative or absolute retention times - ion ratios - signal to noise ratios - lack of interference by chlorinated diphenyl ethers
Calculations	- spot checks of calculations for accuracy
Samples	- results (i.e., concentrations, qualitative identification, and quantitation) assessed with reference to all subjective validation elements

3.0 Supporting Data Elements and Criteria

The review and assessment of supporting data elements is an important part of the validation process. The supporting data elements are assessed for compliance with the appropriate standard, which may be the laboratory's project-specific SOW, reference methods, or Navy's chemical data QA guidance document. The effects of noncompliance on data validity, usability, and defensibility must be assessed, and data flagged as necessary. It is noted that assessment of these elements does always result in technical qualification of data, but may have significant impact on data usability, and technical acceptability. Table H-3 summarizes supporting data elements that are reviewed during the validation process:

Table H-3. Supporting Data Elements Reviewed During Data Validation

Validation Element	Criteria and standards to be assessed during validation
Narrative	- relevant information, e.g., date of sample receipt, date(s) of sample preparation, date(s) of sample analyses, sample matrix, results, and dilution factors - QC failures - initiation of corrective actions - basis of wet or dry weight reporting
Type and Frequency of QC Samples	- compliance with requirements of each reference method and other source documents
Standard Material Traceability and Quality	- traceable to manufacturer and lot number - within assigned and appropriate shelf lives - preparation and use of intermediate and working standards - unique, unambiguous identification of standard materials - preparation of all standard material documented
Reagent traceability and Quality	- traceable to manufacturer and lot number - within assigned shelf life
Analyte List	- complete and accurate target analytes
MDLs/RLs	- frequency of generation - technical acceptability of MDLs - reasonableness of MDLs/RLs - relationship to reporting limit or project required limit(s)
Sample Receipt Conditions	- cooler and individual sample container integrity - temperature - preservation - appropriate containers for analytes - head space

Chain of Custody (CoC)	- unbroken custody record from date and time of sampling through all analyses - internal custody control for extracts and digestates
Sample Storage Conditions	- temperature - preservation - segregation - first removal of volatile aliquots
Unique Identification for Individual Samples	- chain of custody (CoC) - preparative and determinative logs
Dilutions	- documentation - calculation algorithms - appropriate calibration range - correct reporting of results - diluted/undiluted results comparison
Batching Protocol	- batching practices for digestion/extraction and analysis - correlation of samples with associated QC samples and standards.
Pipette Verification	- pipet ID numbers documented - daily calibration check records
Support Equipment (e.g., pH meter, balance, ovens)	- calibration records - daily QC check records - traceability of reference materials and equipment
Corrections/Manual edits	- complete documentation
Corrective Actions	- documentation of nonconformances - nonconformances discussed in the narrative - evaluate laboratory's assessment regarding data quality and usability, if presented
Preparative/Analytical Method	- methods appropriate to sample matrix, analytes, and project requirements - methods can achieve required project limits - method version consistently and accurately specified
Percent Solids	- determined using appropriate protocol - samples reported on dry weight basis if required
Data Review	- scope and levels of review documented
Data Qualifiers	- applicability of data qualifiers assigned by the laboratory - qualifiers defined

Preparation Logs and Run Logs	- completeness - accuracy - analyst and reviewer signatures and dates
Instrument Printouts	- completeness - accuracy - analyst signature and date
Calculations	- calculations spot checked for accuracy
Samples	- results assessed with reference to the supporting elements listed above